State estimation in quantum homodyne tomography with noisy data

J M Aubry[1], C Butucea[2] and K Meziani[3]

[1] Laboratoire d'Analyse et de Mathmatiques Appliques (UMR CNRS 8050),

Universit Paris-Est, 94010 Crteil Cedex, France

[2] Laboratoire Paul Painlev (UMR CNRS 8524),

Universit des Sciences et Technologies de Lille 1, 59655 Villeneuve d'Ascq Cedex, France

[3] Laboratoire de Probabilits et Modles Alatoires,

Universit Paris VII (Denis Diderot), 75251 Paris Cedex 05, France

email: jmaubry@math.cnrs.fr; cristina.butucea@math.univ-lille1.fr; meziani@math.jussieu.fr

26 juillet 2008

Résumé

In the framework of noisy quantum homodyne tomography with efficiency parameter $0 < \eta \leq 1$, we propose two estimators of a quantum state whose density matrix elements $\rho_{m,n}$ decrease like $e^{-B(m+n)^{r/2}}$, for fixed known $B > 0$ and $0 < r \leq 2$. The first procedure estimates the matrix coefficients by a projection method on the pattern functions (that we introduce here for $0 < \eta \leq 1/2$), the second procedure is a kernel estimator of the associated Wigner function. We compute the convergence rates of these estimators, in $\mathbb{L}_2$ risk.

Keywords : density matrix, Gaussian noise, $\mathbb{L}_2$-risk, nonparametric estimation, pattern functions, quantum homodyne tomography, quantum state, Radon transform, Wigner function.

AMS 2000 subject classifications : 62G05, 62G20, 81V80

1 Introduction

Experiments in quantum optics consist in creating, manipulating and measuring quantum states of light. The technique called quantum homodyne tomography allows to retrieve partial, noisy information from which the state is to be recovered : this is the subject of the present chapter.

1.1 Quantum states

Mathematically, the main concepts of quantum mechanics are formulated in the language of selfadjoint operators acting on Hilbert spaces. To every quantum system one can associate a complex Hilbert space $\mathcal{H}$ whose vectors represent the wave functions of the system. These vectors are identified to projection operators, or pure states. In general, a state is a mixture of pure states described by a compact operator ρ on $\mathcal{H}$ having the following properties :

1. Selfadjoint : $\rho = \rho^*$, where ρ^* is the adjoint of ρ.
2. Positive : $\rho \geq 0$, or equivalently $\langle \psi, \rho\psi \rangle \geq 0$ for all $\psi \in \mathcal{H}$.
3. Trace one : $\mathrm{tr}(\rho) = 1$.

When $\mathcal{H}$ is separable, endowed with a countable orthonormal basis, the operator ρ is identified to a *density matrix* $[\rho_{m,n}]_{m,n\in\mathbb{N}}$.

The positivity property implies that all the eigenvalues of ρ are nonegative and by the trace property, they sum up to one. In the case of the finite dimensional Hilbert space $\mathbb{C}^d$, the density matrix is simply a positive semi-definite $d \times d$ matrix of trace one. Our setup from now on will be $\mathcal{H} = L^2(\mathbb{R})$, in which case we employ the orthonormal Fock basis made of the Hermite functions

$$h_m(x) := (2^m m! \sqrt{\pi})^{-\frac{1}{2}} H_m(x) e^{-\frac{x^2}{2}} \tag{1}$$

where $H_m(x) := (-1)^m e^{x^2} \frac{d^m}{dx^m} e^{-x^2}$ is the m-th Hermite polynomial. Generalizations to higher dimensions are straightforward.

To each state ρ corresponds a *Wigner distribution* W_ρ, which is defined via its Fourier transform in the way indicated by equation (2) :

$$\widetilde{W_\rho}(u,v) := \iint e^{-i(uq+vp)} W_\rho(q,p)\,dq\,dp := \mathrm{Tr}\big(\rho \exp(-iu\mathbf{Q} - iv\mathbf{P})\big) \tag{2}$$

where $\mathbf{Q}$ and $\mathbf{P}$ are canonically conjugate observables (e.g. electric and magnetic fields) satisfying the commutation relation $[\mathbf{Q}, \mathbf{P}] = i$ (we assume a choice of units such that $\hbar = 1$). It is easily checked that W_ρ is real-valued, has integral $\iint_{\mathbb{R}^2} W_\rho(q,p)dqdp = 1$ and uniform bound $|W_\rho(q,p)| \leq \frac{1}{\pi}$.

For any $\phi \in \mathbb{R}$, the Wigner distribution allows one to easily recover the probability density $x \mapsto p_\rho(x,\phi)$ of $\mathbf{Q}\cos\phi + \mathbf{P}\sin\phi$ by

$$p_\rho(x,\phi) = \mathcal{R}[W_\rho](x,\phi), \tag{3}$$

where $\mathcal{R}$ is the Radon transform defined in equation (4)

$$\mathcal{R}[W_\rho](x,\phi) = \int_{-\infty}^{\infty} W_\rho(x\cos\phi - t\sin\phi,\, x\sin\phi + t\cos\phi)dt. \tag{4}$$

Moreover, the correspondence between ρ and W_ρ is one to one and isometric with respect to the $\mathbb{L}_2$ norms as in equation (5) :

$$\|W_\rho\|_2^2 := \iint |W_\rho(q,p)|^2 dqdp = \frac{1}{2\pi}\|\rho\|_2^2 := \frac{1}{2\pi}\sum_{j,k=0}^{\infty} |\rho_{jk}|^2. \tag{5}$$

From now on we denote by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$ the usual Euclidian scalar product and norm, while $C(\cdot)$ will denote positive constants depending on parameters given in the parentheses.

We suppose that the unknown state belongs to the class $\mathcal{R}(B,r)$ for $B > 0$ and $0 < r \leq 2$ defined by

$$\mathcal{R}(B,r) := \{\rho \text{ quantum state} : |\rho_{m,n}| \leq \exp(-B(m+n)^{r/2})\}. \tag{6}$$

For simplicity, we have chosen to express the results relative to a class which is the intersection of the (positive) ball of radius 1 in some Banach space with the hyperplane $\mathrm{tr}(\rho) = 1$. Another radius for the class would only change the constant C in front of the asymptotic rates of convergence that we will find.

As it will be made precise in Propositions 1 and 2, quantum states in the class given in (6) have fast decreasing and very smooth Wigner functions. From the physical point of view, the choice of such a class of Wigner functions seems to be quite reasonable considering that typical states ρ prepared in the laboratory do satisfy this type of condition.

1.2 Statistical model

Let us describe the statistical model. Consider $(X_1, \Phi_1), \ldots, (X_n, \Phi_n)$ independent identically distributed random variables with values in $\mathbb{R} \times [0, \pi]$ and distribution P_ρ having density $p_\rho(x, \phi)$ (given by (3) with respect to $\frac{1}{\pi}\lambda$, λ being the Lebesgue measure on $\mathbb{R} \times [0, \pi]$. The aim is to recover the density matrix ρ and the Wigner function W_ρ from the observations.

However, there is a slight complication. What we observe are not the variables (X_ℓ, Φ_ℓ) but the noisy ones (Y_ℓ, Φ_ℓ), where

$$Y_\ell := \sqrt{\eta} X_\ell + \sqrt{(1-\eta)/2}\ \xi_\ell, \tag{7}$$

with ξ_ℓ a sequence of independent identically distributed standard Gaussians which are independent of all (X_j, Φ_j). The detection efficiency parameter $0 < \eta \leq 1$ is known from the calibration of the apparatus and we denote by N^η the centered Gaussian density of variance $(1-\eta)/2$, and $\widetilde{N}^\eta$ its Fourier transform. Then the density p_ρ^η of (Y_ℓ, Φ_ℓ) is given by the convolution of the density $p_\rho(\cdot/\sqrt{\eta}, \phi)/\sqrt{\eta}$ with N^η

$$\begin{aligned} p_\rho^\eta(y, \phi) &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\eta}} p_\rho\left(\frac{y-x}{\sqrt{\eta}}, \phi\right) N^\eta(x) dx \\ &=: \left(\frac{1}{\sqrt{\eta}} p_\rho\left(\frac{\cdot}{\sqrt{\eta}}, \phi\right) * N^\eta\right)(y). \end{aligned}$$

In the Fourier domain this relation becomes

$$\mathcal{F}_1[p_\rho^\eta(\cdot, \phi)](t) = \mathcal{F}_1[p_\rho(\cdot, \phi)](t\sqrt{\eta}) \widetilde{N}^\eta(t), \tag{8}$$

where $\mathcal{F}_1$ denotes the Fourier transform with respect to the first variable.

The theoretical foundation of quantum homodyne tomography was outlined in [29] and has inspired the first experiments determining the quantum state of a light field, initially with optical pulses in [26, 25, 19]. The reconstruction of the density from averages of data has been discussed or studied in [10, 9, 20, 1] for $\eta = 1$ (no photon loss). Max-likelihood methods have been studied in [3, 1, 12, 15] and procedure using adaptive tomographic kernels to minimize the variance has been proposed in [11]. The estimation of the density matrix of a quantum state of light in case of efficiency parameter $\frac{1}{2} < \eta \leq 1$ has been discussed in [7, 12, 8] and considered in [23] via the pattern functions for the diagonal elements.

1.3 Outline of the results

The goal of this chapter is to define estimators of both the density matrix and the Wigner function and to compare their performance in $\mathbb{L}_2$ risk. In order to compute estimation risks and to tune the underlying parameters, we define a realistic class of quantum states $\mathcal{R}(B, r)$, depending on parameters $B > 0$ and $0 < r \leq 2$, in which the elements of the density matrix decrease rapidly.

In Section 2, we prove that the fast decay of the elements of the density matrix implies both rapid decay of the Wigner function and of its Fourier transform, allowing us to translate the classes $\mathcal{R}(B, r)$ in terms of Wigner functions.

In Section 3, we give estimators of the density matrix ρ. The legend was somehow forged that no estimation of the matrix is possible when $0 < \eta \leq 1/2$. The physicists argue that their machines actually have high detection efficiency, around 0.8 ; it is nevertheless satisfying to be able to solve this problem in any noise condition. We give here the so-called *pattern functions* to use for estimating the density matrix in the noisy case with *any* value of η between 0 and 1. These pattern functions allow us to solve an inverse problem which becomes (severly) ill-posed when $0 < \eta \leq 1/2$. In this case, we regularize the inverse problem and this introduces a smoothing parameter which we will choose in an optimal way. We compute the upper bounds for the rates achieved by our methods, with $\mathbb{L}_2$ risk measure.

In Section 4, we study a kernel estimator of the Wigner function in $\mathbb{L}_2$ risk, over the same class of Wigner functions. It is a truncated version of the estimator in [4] and tuned accordingly. We compute upper bounds for the rates of convergence of this estimator in $\mathbb{L}_2$ risk.

To conclude, we may infer that the performances of both estimators are comparable. We obtain nearly polynomial rates for the case $r = 2$ and intermediate rates for $0 < r < 2$ (faster than any logarithm, but slower than any polynomial). It is convenient to have methods to estimate directly both representations of a quantum state. The estimator of the matrix ρ can be more easily projected on the space of proper quantum states. On the other hand, we may capture some features of the quantum states more easily on the Wigner function, for instance when this function has significant negative parts, the fact that the quantum state is non classical.

2 Decrease and smoothness of the Wigner distribution

We recall that the Wigner distribution W_ρ was defined in the introduction. In the Fock basis, we can write W_ρ in terms of the density matrix $[\rho_{m,n}]$ as follows (see Leonhardt [19] for the details).

$$W_\rho(q,p) = \sum_{m,n} \rho_{m,n} W_{m,n}(q,p)$$

where

$$W_{m,n}(q,p) = \frac{1}{\pi} \int e^{2ipx} h_m(q-x) h_n(q+x) dx. \tag{9}$$

It can be seen that $W_{m,n}(q,p) = W_{n,m}(q,-p)$ and if $m \geq n$,

$$\begin{aligned} W_{m,n}(q,p) \quad &= \quad \frac{(-1)^m}{\pi} \left(\frac{n!}{m!}\right)^{\frac{1}{2}} e^{-(q^2+p^2)} \\ &\quad \times \left(\sqrt{2}(ip-q)\right)^{m-n} L_n^{m-n}\left(2q^2+2p^2\right) \end{aligned} \tag{10}$$

thus, writing $z := \sqrt{q^2+p^2}$,

$$l_{m,n}(z) := |W_{m,n}(q,p)| = \frac{2^{\frac{m-n}{2}}}{\pi} \left(\frac{n!}{m!}\right)^{\frac{1}{2}} e^{-z^2} z^{m-n} \left| L_n^{m-n}(2z^2) \right| \tag{11}$$

where $L_n^\alpha(x) := (n!)^{-1} e^x x^{-\alpha} \frac{d^n}{dx^n}(e^{-x} x^{n+\alpha})$ is the Laguerre polynomial of degree n and order α. Concerning the Fourier transforms, we also recall that

$$\widetilde{W_{m,n}}(q,p) = \frac{(-i)^{m+n}}{2} W_{m,n}\left(\frac{q}{2}, \frac{p}{2}\right). \tag{12}$$

In this section we show how a decrease condition on the coefficients of the density matrix translates on the corresponding Wigner distribution. First the case $r < 2$:

Proposition 1. *Assume that $0 < r < 2$ and that there exists $B > 0$ such that, for all $m \geq n$,*

$$|\rho_{m,n}| \leq e^{-B(m+n)^{r/2}}.$$

Then for all $\beta < B$, there exists z_0 (depending explicitly on r, B, β, see proof) such that $z := \sqrt{q^2+p^2} \geq z_0$ implies

$$|W_\rho(q,p)| \leq A(z) e^{-\beta z^r} \tag{13}$$

as well as

$$\left| \widetilde{W_\rho}(q,p) \right| \leq A(z/2) e^{-\beta (z/2)^r} \tag{14}$$

where $A(z) := \frac{1}{\pi}\left(\sum_{m,n} e^{-B(m+n)^{r/2}} + \frac{4}{Br} z^{4-r}\right)$.

If $r = 2$, the result is a little different :

Proposition 2. *Suppose that there exists* $B > 0$ *such that, for all* $m \geq n$,

$$|\rho_{m,n}| \leq e^{-B(m+n)}.$$

Then there exists z_0 *such that* $z := \sqrt{q^2 + p^2} \geq z_0$ *implies*

$$|W_\rho(q,p)| \leq A(z) e^{-\frac{B}{(1+\sqrt{B})^2} z^2} \tag{15}$$

as well as

$$\left|\widetilde{W_\rho}(q,p)\right| \leq A(z/2) e^{-\frac{B}{(1+\sqrt{B})^2}(z/2)^2} \tag{16}$$

for $A(z) = \frac{1}{\pi}\left(\sum_{m,n} e^{-B(m+n)} + \frac{2e^B}{B(1+\sqrt{B})^2} z^2\right)$.

Note that $\frac{B}{\left(1+\sqrt{B}\right)^2} < \min(B,1)$. Even when B is very large, we cannot hope to obtain a faster decrease because e^{-z^2} is the decrease rate of the basis functions themselves (Lemma 2).

The proof of these propositions is defered to Appendix 5. More general results and converses are studied in [2]. Let us now state a few general utility lemmata.

Lemma 1. *Let* y *and* w *be two* C^2 *functions* : $[x_0, +\infty) \to (0, +\infty)$ *such that* $y'(x) \to 0$, w *is bounded, satisfying the differential equations*

$$\begin{aligned} y''(x) &= \phi(x) y(x) \\ w''(x) &= \psi(x) w(x), \end{aligned}$$

with continuous $\phi(x) \leq \psi(x)$, *and initial conditions* $y(x_0) = w(x_0)$. *Then for all* $x \geq x_0$, $w(x) \leq y(x)$.

Démonstration. Suppose that there exists $x_1 \geq x_0$ where $w(x_1) > y(x_1)$. Then for some $x_2 \in [x_0, x_1]$ we have $w'(x_2) > y'(x_2)$ and $w(x_2) \geq y(x_2)$. Consequently, for all $x \geq x_2$, $w''(x) - y''(x) \geq 0$, and $w'(x) - y'(x) \geq w'(x_2) - y'(x_2)$. When $x \to \infty$, $\liminf w'(x) \geq w'(x_2) - y'(x_2) > 0$, which contradicts the boundedness of w. □

This lemma is used to prove a bound on the Laguerre functions.

Lemma 2. *For all $m, n \in \mathbb{N}$ and $s := \sqrt{m+n+1}$, for all $z \geq 0$,*

$$l_{m,n}(z) \leq \frac{1}{\pi} \begin{cases} 1 & \text{if } 0 \leq z \leq s \\ e^{-(z-s)^2} & \text{if } z \geq s. \end{cases} \tag{17}$$

Démonstration. When $z \leq s$, the result follows from the uniform bound on Wigner functions obtained by applying the Cauchy-Schwarz inequality to (9).
When $z \geq s$, $L_n^\alpha(2z^2)$ doesn't vanish and keeps the same sign as $L_n^\alpha(2s^2)$. Now, as it can be seen from [27, 5.1.2], the function $w(z) := \sqrt{z} l_{m,n}(z)$ satisfies the differential equation $w'' = (4(z^2 - s^2) + \frac{\alpha^2 - 1/4}{z^2})z$. On the other hand, $y(z) := \sqrt{s} l_{m,n}(s) e^{-(z-s)^2}$ satisfies $y'' = (4(z-s)^2 - 2)y$. When $z \geq s$,

$$4(z-s)^2 - 2 < 4(z^2 - s^2) + \frac{\alpha^2 - 1/4}{z^2} \tag{18}$$

from which we conclude with Lemma 1 that $w(z) \leq y(z)$. □

Finally, a lemma to bound the tail of a series.

Lemma 3. *If $\nu > 0$ and $C > 0$, there exists a z_0 such that $z \geq z_0$ implies*

$$\sum_{m+n \geq z} e^{-C(m+n)^\nu} \leq \frac{2}{C\nu} z^{2-\nu} e^{-Cz^\nu}. \tag{19}$$

Démonstration. First notice that

$$\sum_{m+n \geq z} e^{-C(m+n)^\nu} = \sum_{t \geq z} (t+1) e^{-Ct^\nu} \leq \int_z^\infty (t+1) e^{-Ct^\nu} dt.$$

When $t \geq z$ and z is large enough, we have

$$\begin{aligned} \int_z^\infty (t+1) e^{-Ct^\nu} dt &\leq \frac{2}{C\nu} \int_z^\infty \big(C\nu t - (2-\nu) t^{1-\nu}\big) e^{-Ct^\nu} dt \\ &\leq \frac{2}{C\nu} z^{2-\nu} e^{-Cz^\nu} \end{aligned}$$

which is what we needed to prove. □

3 Density matrix estimation

The aim of this part is to estimate the density matrix ρ in the Fock basis directely from the data $(Y_i, \Phi_i)_{i=1,\dots,n}$. We show that for $0 < \eta \leq 1/2$ it is still possible to estimate the density matrix with an error of estimation tending to 0 as n tends to infinity (Theorem 3). In both cases ($\eta > \frac{1}{2}$ and $\eta \leq \frac{1}{2}$), we construct an estimator of the density matrix $(\rho_{j,k})_{j,k\leq N-1}$ from a sample of QHT data. We give theoretical results for our estimator when the quantum state ρ is in the class of density matrix with decreasing elements defined in (6).

3.1 Pattern functions

The matrix elements $\rho_{j,k}$ of the state ρ in the Fock basis (1) can be expressed as kernel integrals : for all $j, k \in \mathbb{N}$,

$$\rho_{j,k} = \frac{1}{\pi} \int \int_0^\pi p_\rho(x, \phi) f_{j,k}(x) e^{-i(k-j)\phi} d\phi dx \tag{20}$$

where $f_{j,k} = f_{k,j}$ are bounded real functions called *pattern functions* in quantum homodyne literature. A concrete expression for their Fourier transform using Laguerre polynomials was found in [24] : for $j \geq k$,

$$\begin{aligned} \tilde{f}_{k,j}(t) &= 2\pi^2 |t| \widetilde{W_{j,k}}(t, 0) \\ &= \pi(-i)^{j-k} \sqrt{\frac{2^{k-j} k!}{j!}} |t| t^{j-k} e^{-\frac{t^2}{4}} L_k^{j-k}(\frac{t^2}{2}). \end{aligned} \tag{21}$$

where $\tilde{f}_{k,j}$ denotes the Fourier transform of the Pattern function $f_{k,j}$.
Let us state the lemmata which are used to prove upper bounds in Propositions 3, 4 and 5.

Lemma 4. *There exist constants C_2, C_∞ such that*

$$\textstyle\sum_{j+k=0}^N \|f_{k,j}\|_2^2 \leq C_2 N^{\frac{17}{6}} \ \textit{and} \quad \sum_{j+k=0}^N \|f_{k,j}\|_\infty^2 \leq C_\infty N^{\frac{10}{3}}.$$

This is a slight improvement over [1, Lemma 1].

Démonstration. By symmetry we can restrict the sum to $j \geq k$. For fixed k and j we have

$$\left\| \tilde{f}_{k,j} \right\|_2^2 = \int_{|t|<2s} \left| \tilde{f}_{k,j}(t) \right|^2 dt + \int_{|t|>2s} \left| \tilde{f}_{k,j}(t) \right|^2 dt$$

(with $s = \sqrt{k+j+1}$). Because of Lemma 2, it is clear that the second integral is negligible in front of the first one, which we simply bound by $4s \left\| \tilde{f}_{k,j} \right\|_\infty^2$.
In view of (21), the main result in [18] can be rewritten as follows : if $k \geq 35$ and $j - k \geq 24$, then

$$\left\| \tilde{f}_{k,j} \right\|_\infty^2 \leq 2888\pi^2 (j+1)^{\frac{1}{2}} k^{-\frac{1}{6}}. \tag{22}$$

In consequence, for these values of k and j,

$$\left\|\tilde{f}_{k,j}\right\|_2^2 \leq C(jk^{-\frac{1}{6}} + j^{\frac{1}{2}}k^{\frac{1}{3}}). \tag{23}$$

On the other hand, a classical bound on Laguerre polynomials found in [27] yields that, for fixed values of $j-k$, $\left\|\tilde{f}_{k,j}\right\|_\infty^2 \leq Ck^{\frac{1}{3}}$, hence for all $k \geq 35$ and $j-k < 24$,

$$\left\|\tilde{f}_{k,j}\right\|_2^2 \leq C(j^{\frac{1}{2}}k^{\frac{1}{3}} + k^{\frac{5}{6}}). \tag{24}$$

When $k < 35$, we can use another result in [17] which gives $\left\|\tilde{f}_{k,j}\right\|_\infty^2 \leq Ck^{\frac{1}{6}}j^{\frac{1}{2}}$ independently of $j-k$, thus

$$\left\|\tilde{f}_{k,j}\right\|_2^2 \leq Cj. \tag{25}$$

Comparing (23), (24) and (25) we see that when N is large enough, in the sum over $0 \leq j,k \leq N$, the terms $k \geq 35$, $j-k \geq 24$ dominate and (23) yields the first inequality. The second inequality is obtained by doing a similar computation, starting with $\|f_{j,k}\|_\infty \leq \left\|\tilde{f}_{j,k}\right\|_1$ and using (22) to bound

$$\left\|\tilde{f}_{j,k}\right\|_1^2 \leq C(j^{\frac{3}{2}}k^{-\frac{1}{6}} + j^{\frac{1}{2}}k^{\frac{5}{6}})$$

when $k \geq 35$ and $j-k \geq 24$. □

In the presence of noise, it is necessary to adapt the pattern functions as follows. From now on, we shall use the notation $\boxed{\gamma := \frac{1-\eta}{4\eta}}$. When $\frac{1}{2} < \eta \leq 1$, we denote by $f_{k,j}^\eta$ the function which has the following Fourier transform :

$$\tilde{f}_{k,j}^\eta(t) := \tilde{f}_{k,j}(t)e^{\gamma t^2}. \tag{26}$$

When $0 < \eta \leq \frac{1}{2}$, we introduce a cut-off parameter $\delta > 0$ and define $f_{k,j}^{\eta,\delta}$ via its Fourier transform :

$$\tilde{f}_{k,j}^{\eta,\delta}(t) := \tilde{f}_{k,j}(t)e^{\gamma t^2}\mathbb{I}\left(|t| \leq \frac{1}{\delta}\right). \tag{27}$$

Then we compute bounds on these pattern functions.

Lemma 5. *For $1 > \eta > 1/2$, there exist constants C_2^η and C_∞^η such that*

$$\sum_{j+k=0}^N \left\|f_{k,j}^\eta\right\|_2^2 \leq C_2^\eta N^{\frac{5}{6}}e^{8\gamma N} \quad \text{and} \quad \sum_{j+k=0}^N \left\|f_{k,j}^\eta\right\|_\infty^2 \leq C_\infty^\eta N^{\frac{1}{3}}e^{8\gamma N}.$$

Démonstration. The proof is similar to the previous one and we skip some details. Once again we assume $j \geq k$ and write

$$\left\|\tilde{f}^{\eta}_{k,j}\right\|_2^2 = \int_{|t|<2s} \left|\tilde{f}_{k,j}(t)\right|^2 e^{2\gamma t^2} dt + \int_{|t|>2s} \left|\tilde{f}_{k,j}(t)\right|^2 e^{2\gamma t^2} dt$$

(where $s = \sqrt{k+j+1}$). Because of Lemma 2, the second integral is of the same order as the first one, which we bound by

$$\left\|\tilde{f}_{k,j}\right\|_\infty^2 \int_{|t|<2s} e^{2\gamma t^2} dt \leq C \left\|\tilde{f}_{k,j}\right\|_\infty^2 s^{-1} e^{8\gamma s^2}.$$

In the sum we are considering the terms $k \geq 35$ and $j - k \geq 24$ are dominant and, once again thanks to (22), remembering that $s = \sqrt{j+k+1}$,

$$\left\|\tilde{f}^{\eta}_{k,j}\right\|_2^2 \leq C k^{-\frac{1}{6}} e^{8\gamma(j+k)}$$

hence the first inequality.

The second inequality is, in the same fashion, based on

$$\begin{aligned} \left\|f^{\eta}_{k,j}\right\|_\infty^2 \leq \left\|\tilde{f}^{\eta}_{k,j}\right\|_1^2 &\leq C\left(j^{\frac{1}{4}} k^{-\frac{1}{12}} \int_{|t|<2s} e^{\gamma t^2} dt\right)^2 \\ &\leq C j^{-\frac{1}{2}} k^{-\frac{1}{6}} e^{8\gamma(j+k)} \end{aligned}$$

when $k \geq 35$ and $j - k \geq 24$, and the bound on the sum readily follows. □

3.2 Estimation procedure

For $N := N(n) \to \infty$ and $\delta := \delta(n) \to 0$, let us define our estimator of $\rho_{j,k}$ for $0 \leq j+k \leq N-1$ by

$$\hat{\rho}^{\eta}_{j,k} := \frac{1}{n} \sum_{\ell=1}^{n} G_{j,k}\left(\frac{Y_\ell}{\sqrt{\eta}}, \Phi_\ell\right), \tag{28}$$

where

$$G_{j,k}(x,\phi) := \begin{cases} f^{\eta}_{j,k}(x) e^{-i(j-k)\phi} & \text{if } \frac{1}{2} < \eta \leq 1 \\ f^{\eta,\delta}_{j,k}(x) e^{-i(j-k)\phi} & \text{if } 0 < \eta \leq \frac{1}{2}. \end{cases}$$

using the pattern functions defined in (26) and (27). We assume that the density matrix ρ belongs to the class $\mathcal{R}(B,r)$ defined in (6). In order to evaluate the performance of our estimators we take the $\mathbb{L}_2$ distance on the space of density matrices $\|\tau - \rho\|_2^2 := \mathrm{tr}(|\tau -$

$\rho|^2) = \sum_{j,k=0}^{\infty} |\tau_{j,k} - \rho_{j,k}|^2$. We consider the mean integrated square error (MISE) and split it into a troncature bias term $b_1^2(n)$, a regularization bias terms $b_2^2(n)$ and a variance term $\sigma^2(n)$.

$$\begin{aligned} E\left(\sum_{j,k=0}^{\infty} \left|\hat{\rho}_{j,k}^{\eta} - \rho_{j,k}\right|^2\right) &= \sum_{j+k\geq N} |\rho_{j,k}|^2 + \sum_{j+k=0}^{N-1} \left|E[\hat{\rho}_{j,k}^{\eta}] - \rho_{j,k}\right|^2 \\ &\quad + \sum_{j+k=0}^{N-1} E\left|\hat{\rho}_{j,k}^{\eta} - E[\hat{\rho}_{j,k}^{\eta}]\right|^2 \\ &=: b_1^2(n) + b_2^2(n) + \sigma^2(n). \end{aligned}$$

The following propositions give upper bounds for $b_1^2(n)$, $b_2^2(n)$ and $\sigma^2(n)$ in the different cases $\eta = 1$, $1/2 < \eta < 1$ or $0 < \eta \leq 1/2$ and $r = 2$ or $0 < r < 2$. Their proofs are defered to Appendix 5.

Proposition 3. *Let $\hat{\rho}_{j,k}^{\eta}$ be the estimator defined by (28), for $0 < \eta < 1$, with $\delta \to 0$ and $N \to \infty$ as $n \to \infty$, then for all $B > 0$ and $0 < r \leq 2$,*

$$\sup_{\rho\in\mathcal{R}(B,r)} b_1^2(n) \leq c_1 N^{2-r/2} e^{-2BN^{r/2}} \tag{29}$$

where c_1 is a positive constant depending on B and r.

Proposition 4. *Let $\hat{\rho}_{j,k}^{\eta}$ be the estimator defined by (28), for $0 < \eta \leq 1/2$, with $N \to \infty$ as $n \to \infty$ and $1/\delta \geq 2\sqrt{N}$. In the case $r = 2$, for $\beta := B/(1+\sqrt{B})^2$ there exists c_2, while in the case $0 < r < 2$, for any $\beta < B$ there exists c_2 and n_0 such that for $n \geq n_0$:*

$$\sup_{\rho\in\mathcal{R}(B,r)} b_2^2(n) \leq c_2 N^2 \delta^{4r-12} e^{-\frac{2\beta}{(2\delta)^r} - \frac{1}{2}\left(\frac{1}{\delta} - 2\sqrt{N}\right)^2}. \tag{30}$$

Note that for $1/2 < \eta \leq 1$ we have $b_2(n) = 0$ for all $0 < r \leq 2$ ($\hat{\rho}_{j,k}^{\eta}$ is unbiased).

Proposition 5. *For $\hat{\rho}_{j,k}^{\eta}$ the estimator defined by (28),*

$$\sup_{\rho\in\mathcal{R}(B,r)} \sigma^2(n) \leq c_3 \frac{\delta N^{17/6}}{n} e^{\frac{2\gamma}{\delta^2}} \qquad \text{if } 0 < \eta \leq 1/2 \tag{31}$$

$$\sup_{\rho\in\mathcal{R}(B,r)} \sigma^2(n) \leq c_3' \frac{N^{1/3}}{n} e^{8\gamma N} \qquad \text{if } 1/2 < \eta < 1 \tag{32}$$

$$\sup_{\rho\in\mathcal{R}(B,r)} \sigma^2(n) \leq c_3'' \frac{N^{\frac{17}{6}}}{n} \qquad \text{if } \eta = 1 \tag{33}$$

where c_3, c_3' are positive constants depending on η.

We measure the accuracy of $\hat{\rho}^{\eta}_{j,k}$ by the maximal risk over the class $\mathcal{R}(B,r)$

$$\limsup_{n\to\infty} \sup_{\rho\in\mathcal{R}(B,r)} \varphi_n^{-2} E\left(\sum_{j,k=0}^{\infty} \left|\hat{\rho}^{\eta}_{j,k} - \rho_{j,k}\right|^2\right) \leq C_0. \tag{34}$$

where C_0 is a positive constant and φ_n^2 is a sequence which tends to 0 when $n \to \infty$ and it is the rate of convergence. Cases $\eta = 1$ (no noise), $\frac{1}{2} < \eta < 1$ (weak noise) and $0 < \eta \leq \frac{1}{2}$ (strong noise) are studied respectively in Theorems 1, 2 and 3.

Theorem 1. *When $\eta = 1$, the estimator defined in (28) for the model (7), where the unknown state belongs to the class $\mathcal{R}(B,r)$, satisfies the upper bound (34) with*

$$\varphi_n^2 = \log(n)^{\frac{17}{3r}} n^{-1}$$

obtained by taking $N(n) := \left(\frac{\log(n)}{2B}\right)^{\frac{2}{r}}$.

Démonstration. With the proposed $N(n)$ one checks that the bias (29) is smaller than the variance (33) which is bounded by a constant times $\log(n)^{\frac{17}{3r}} n^{-1}$. □

Theorem 2. *When $\frac{1}{2} < \eta < 1$, the estimator defined in (28) for the model (7), where the unknown state belongs to the class $\mathcal{R}(B,r)$, satisfies the upper bound (34) with*

– *For $r = 2$,*

$$\varphi_n^2 = \log(n)^{\frac{12\gamma+B}{3(4\gamma+B)}} n^{-\frac{B}{4\gamma+B}}$$

with $N(n) := \frac{\log(n)}{2(4\gamma+B)}\left(1 + \frac{2}{3}\frac{\log(\log n)}{\log(n)}\right)$.

– *For $0 < r < 2$,*

$$\varphi_n^2 = \log(n)^{2-r/2} e^{-2BN(n)^{r/2}}$$

where $N(n)$ is the solution of the equation $8\gamma N + 2BN^{r/2} = \log(n)$.

In that case we have $N(n) = \frac{1}{8\gamma}\log(n) - \frac{2B}{(8\gamma)^{1+r/2}}\log(n)^{r/2} + o(\log(n)^{r/2})$.

Démonstration. When $r = 2$, the proposed $N(n)$ ensures that the variance (32) is equivalent to the bias (29), which is bounded by a constant times $\log(n)^{\frac{12\gamma+B}{3(4\gamma+B)}} n^{-\frac{B}{4\gamma+B}}$.

When $0 < r < 2$, the proposed $N(n)$ makes the variance (32) bounded by a constant times $e^{-2BN(n)^{r/2}}$, which is smaller than the bias, the latter being bounded by a constant times $N(n)^{2-r/2}e^{-2BN(n)^{r/2}}$.

The asymptotic expansion of $N(n)$ is a standard consequence of its definition by the equation $8\gamma N + 2BN^{r/2} = \log(n)$. □

Theorem 3. *When $0 < \eta \leq \frac{1}{2}$, the estimator defined in (28) for the model (7), where the unknown state belongs to the class $\mathcal{R}(B, r)$, satisfies the upper bound (34) with*

$$\varphi_n^2 = N^{2-r/2} e^{-2BN^{r/2}}$$

where N and δ are solutions of the system

$$\left\{ \begin{array}{l} \frac{2\beta}{(2\delta)^r} + \frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 + \frac{2\gamma}{\delta^2} = \log(n) \\ \frac{2\beta}{(2\delta)^r} + \frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 - 2BN^{r/2} = (\log\log(n))^2 \end{array} \right. \tag{35}$$

for arbitrary $\beta < B$ in the case $0 < r < 2$ or

$$\left\{ \begin{array}{l} \frac{\beta+4\gamma}{2\delta^2} + \frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 - \frac{5}{3}\log(N) = \log(n) \\ \frac{\beta}{2\delta^2} + \frac{1}{2}(\frac{1}{\delta} - 2\sqrt{N})^2 - 2BN - 3\log(N) = 0 \end{array} \right. \tag{36}$$

with $\beta := \frac{B}{(1+\sqrt{B})^2}$ in the case $r = 2$.

Theses bounds are optimal in the sense that (35) and (36) are obtained by minimizing the sum of the bounds (29), (30) and (31).

Démonstration. We use the standard notations $a(n) \sim b(n)$ if $\frac{a(n)}{b(n)} \to 1$ and $a(n) \approx b(n)$ if there exists a constant $M < \infty$ such that $\frac{1}{M} \leq \frac{a(n)}{b(n)} \leq M$ for all n.

Let us first examine the case $0 < r < 2$. Remark that the left-hand term of the second equation in (35) is strictly negative when $1/\delta = 2\sqrt{N}$ and increases to ∞ with $1/\delta$. This proves that the solution satisfies $1/\delta > 2\sqrt{N}$ and that Proposition 4 applies. Furthermore, if we suppose that $\frac{1/\delta}{\sqrt{N}}$ is unbounded when $n \to \infty$, then (up to taking a subsequence) by the first equation $\frac{\frac{1}{2}+2\gamma}{\delta^2} \sim \log(n)$ whereas, by subtracting the two, $\frac{2\gamma}{\delta^2} \sim \log(n)$, which is contradictory. So $1/\delta \approx \sqrt{N}$ and we deduce that $N \approx \log(n)$. Then (30) yields

$$\log\left(\frac{b_2^2(n)}{N^{2-r/2}e^{-2BN^{r/2}}}\right) \leq (4r-12)\log(\delta) + \frac{r}{2}\log(N) - (\log\log(n))^2 \to -\infty$$

whereas (31) gives

$$\log\left(\frac{\sigma^2(n)}{N^{2-r/2}e^{-2BN^{r/2}}}\right) \leq \log(\delta) + (\frac{5}{6} + \frac{r}{2})\log(N) - (\log\log(n))^2 \to -\infty.$$

We see that the dominant term is the bound (29) on $b_1^2(n)$, hence the result.

When $r = 2$, the same reasoning as above yields $1/\delta > 2\sqrt{N}$, $1/\delta \approx \sqrt{N}$ and $N \approx \log(n)$. Then the right-hand side of (30) and (31) are of the same order as Ne^{-2BN}, which is the bound (29) on $b_1^2(n)$. □

4 Wigner function estimation

4.1 Kernel estimator

We describe now the direct estimation method for the Wigner function. For the problem of estimating a probability density $f : \mathbb{R}^2 \to \mathbb{R}$ directly from data (X_ℓ, Φ_ℓ) with density $\mathcal{R}[f]$ we refer to the literature on X-ray tomography and PET, studied by [28, 16, 21, 6] and many other references therein. In the context of tomography of bounded objects with noisy observations [13] solved the problem of estimating the borders of the object (the support). The estimation of a quadratic functional of the Wigner function has been treated in [22]. For the problem of Wigner function estimation when no noise is present, we mention the work by [14]. They use a kernel estimator and compute sharp minimax results over a class of Wigner functions characterised by their smoothness. In a more recent paper [4], Butucea, Guţă and Artiles treated the noisy problem for the pointwise estimation of W_ρ; however the functions needed to prove minimax optimality there do not belong to the class of Wigner functions that we consider here.

In this chapter, as in [4], we modify the usual tomography kernel in order to take into account the additive noise on the observations and construct a kernel K_h^η which performs both deconvolution and inverse Radon transform on our data, asymptotically. Let us define the estimator :

$$\widehat{W}_h^\eta(q,p) = \frac{1}{\pi n}\sum_{\ell=1}^{n} K_h^\eta\left(q\cos\Phi_\ell + p\sin\Phi_\ell - \frac{Y_\ell}{\sqrt{\eta}}\right), \tag{37}$$

where $0 < \eta < 1$ is a fixed parameter, and the kernel is defined by

$$K_h^\eta(u) = \frac{1}{4\pi}\int_{-1/h}^{1/h} \frac{\exp(-iut)|t|}{\widetilde{N}^\eta(t/\sqrt{\eta})}dt, \quad \widetilde{K}_h^\eta(t) = \frac{1}{2}\frac{|t|}{\widetilde{N}^\eta(t/\sqrt{\eta})} I(|t| \le 1/h), \tag{38}$$

and $h > 0$ tends to 0 when $n \to \infty$ in a proper way to be chosen later. For simplicity, let us denote $z = (q,p)$ and $[z,\phi] = q\cos\phi + p\sin\phi$, then the estimator can be written :

$$\widehat{W}_h^\eta(z) = \frac{1}{\pi n}\sum_{\ell=1}^{n} K_h^\eta\left([z,\Phi_\ell] - \frac{Y_\ell}{\sqrt{\eta}}\right).$$

This is a one-step procedure for treating two successive inverse problems. The main difference with the noiseless problem treated by [14] is that the deconvolution is more ‘difficult’ than the inverse Radon transform. In the literature on inverse problems, this problem would be qualified as severely ill-posed, meaning that the noise is dramatically (exponentially) smooth and makes the estimation problem much harder.

4.2 $\mathbb{L}_2$ risk estimation

We establish next the rates of estimation of W_ρ from i.i.d. observations (Y_ℓ, Φ_ℓ), $\ell = 1, \ldots, n$ when the quality of estimation is measured in $\mathbb{L}_2$ distance. In the literature, $\mathbb{L}_2$ tomography is usually performed for boundedly supported functions, see [16] and [21]. However, most Wigner function do not have a bounded support! Instead, we use the fact that Wigner functions in the class $\mathcal{R}(B, r)$ decrease very fast and show that a properly truncated estimator attains the rates we may expect from the statistical problem of deconvolution in presence of tomography. Thus, we modify the estimator by truncating it over a disc with increasing radius, as $n \to \infty$. Let us denote

$$D(s_n) = \{z = (q,p) \in \mathbb{R}_2 : \|z\| \leq s_n\},$$

where $s_n \to \infty$ as $n \to \infty$ will be defined in Theorem 4. Let now

$$\widehat{W}_{h,n}^{\eta,*}(z) = \widehat{W}_{h,n}^{\eta}(z) I_{D(s_n)}(z). \tag{39}$$

From now on, we will denote for any function f,

$$\|f\|_{D(s_n)}^2 = \int_{D(s_n)} f^2(z)dz,$$

and by $\overline{D}(s_n)$ the complementary set of $D(s_n)$ in $\mathbb{R}^2$. Then,

$$\begin{aligned} E\left[\left\|\widehat{W}_h^{\eta,*} - W_\rho\right\|_2^2\right] &= E\left[\left\|\widehat{W}_h^{\eta} - W_\rho\right\|_{D(s_n)}^2\right] + \|W_\rho\|_{\overline{D}(s_n)}^2 \\ &= E\left[\left\|\widehat{W}_h^{\eta} - E\left[\widehat{W}_h^{\eta}\right]\right\|_{D(s_n)}^2\right] + \left\|E\left[\widehat{W}_h^{\eta}\right] - W_\rho\right\|_{D(s_n)}^2 \\ &\quad + \|W_\rho\|_{\overline{D}(s_n)}^2. \end{aligned}$$

When replacing the $\mathbb{L}_2$ norm with the above restricted integral, the upper bound of the bias of the estimator is unchanged, whereas the variance part is infinitely larger than the deconvolution variance in [5]. As the bias is dominating over the variance in this setup, we can still choose a suitable sequence s_n so that the same bandwidth is optimal associated to the same optimal rate, provided that W_ρ decreases fast enough asymptotically. The following proposition gives upper bounds for the three components of the $\mathbb{L}_2$ risk uniformly over the class $\mathcal{R}(B,r)$.

Proposition 6. *Let (Y_ℓ,Φ_ℓ), $\ell=1,\dots,n$ be i.i.d. data coming from the model (7) and let $\widehat{W}_h^\eta$ be an estimator (with $h\to 0$ as $n\to\infty$) of the underlying Wigner function W_ρ. We suppose W_ρ lies in the class $\mathcal{R}(B,r)$, with $B>0$ and $0<r\le 2$. Then, for $s_n\to\infty$ as $n\to\infty$ and n large enough,*

$$\begin{aligned}
\sup_{\rho\in\mathcal{R}(B,r)} \|W^\rho\|^2_{\bar{D}(s_n)} &\le C_1 s_n^{10-3r} e^{-2\beta s_n^r},\\
\sup_{\rho\in\mathcal{R}(B,r)} \left\| E[\widehat{W}_h^\eta] - W_\rho \right\|^2_{D(s_n)} &\le C_2 h^{3r-10} e^{-\frac{2^{1-r}\beta}{h^r}}\\
\sup_{\rho\in\mathcal{R}(B,r)} E\left[\left\| \widehat{W}_{h,n}^\eta - E\left[\widehat{W}_{h,n}^\eta\right]\right\|^2_{D(s_n)}\right] &\le C_3 \frac{s_n^2}{nh}\exp\left(\frac{2\gamma}{h^2}\right),
\end{aligned}$$

where $\beta<B$ is defined in Proposition 1 for $0<r<2$ and $\beta=B/(1+\sqrt{B})^2$ for $r=2$, $\gamma=(1-\eta)/(4\eta)>0$, C_1, C_2, C_3 are positive constants, C_1, C_2, depending on β, B, r and C_3 depending only on η.

We measure the accuracy of $\widehat{W}_h^{\eta,*}$ by the maximal risk over the class $\mathcal{R}(B,r)$

$$\limsup_{n\to\infty} \sup_{\rho\in\mathcal{R}(B,r)} E\left[\left\|\widehat{W}_h^{\eta,*} - W_\rho\right\|^2\right]\varphi_n^{-2}(\mathbb{L}_2) \le C. \tag{40}$$

where C is a positive constant and φ_n^2 is a sequence which tends to 0 when $n\to\infty$ and it is the rate of convergence.

In the following Theorem we see the phenomenon which was noticed already : deconvolution with Gaussian type noise is a much harder problem than inverse Radon transform (the tomography part).

Theorem 4. *Let $B>0$, $0<r\le 2$ and (Y_ℓ,Φ_ℓ), $\ell=1,\dots,n$ be i.i.d. data coming from the model (7). Then $\widehat{W}_h^{\eta,*}$ defined in (39) with kernel K_h^η in (38) satisfies the upper bound (40) with*

- *For $r=2$, put $\beta=B/(1+\sqrt{B})^2$*

$$\varphi_n^2 = (\log n)^{\frac{16\gamma+3\beta}{8\gamma+2\beta}} n^{-\frac{\beta}{4\gamma+\beta}},$$

with $s_n=(h)^{-1}$ and $h=\left(\frac{2}{4\gamma+\beta}\log n + \frac{1}{4\gamma+\beta}\log(\log n)\right)^{-1/2}$.

– *For* $0 < r < 2$ *and* $\beta < B$ *defined in Proposition 1,*

$$\varphi_n^2 = h^{3r-10} \exp\left(-\frac{2^{1-r}\beta}{h^r}\right),$$

where $s_n = 1/h$ *and* h *is the solution of the equation*

$$\frac{2^{1-r}\beta}{h^r} + \frac{2\gamma}{h^2} = \log n - (\log\log n)^2.$$

Sketch of proof of the upper bounds. By Proposition 6, we get

$$\begin{aligned}
\sup_{W_\rho \in \mathcal{R}(B,r)} E\left[\left\|\widehat{W}_h^\eta - W_\rho\right\|^2\right] &\le C_1 s_n^{10-3r} e^{-2\beta s_n^r} + C_2 h^{3r-10} \exp\left(-\frac{2\beta}{(2h)^r}\right) \\
&\quad + \frac{C_3 s_n^2}{nh} \exp\left(\frac{2\gamma}{h^2}\right). \\
&=: A_1 + A_2 + A_3
\end{aligned}$$

For $0 < r < 2$ and by taking derivatives with respect to h and s_n, we obtain that the optimal choice verifies the following equations :

$$\begin{aligned}
2\beta s_n^r + \frac{2\gamma}{h^2} &= \log(n) + \log(h s_n^{2(4-r)}) \\
\frac{2^{1-r}\beta}{h^r} + \frac{2\gamma}{h^2} &= \log(n) + \log(h^{2r-7} s_n^{-2}).
\end{aligned}$$

We notice therefore that A_2 is dominating over A_3, which is dominating over A_1. The proposed (s_n, h) ensure that the term A_2 is still the dominating term and gives the rate of convergence.

The case $r = 2$ is treated similarly, by taking derivatives we notice that the term A_2 and the term A_3 are of the same order and that the term A_1 is smaller than the others. □

5 Appendix

5.1 Proof of Proposition 1

Let $\phi(z) := (z - \sqrt{\beta} z^{r/2})^2 - 1$. Since $r < 2$, for z larger than a certain z_0 (which depends only on β, B and r), it is true that $\phi(z) \ge \left(\frac{\beta}{B}\right)^{2/r} z^2$. It follows that

$$e^{-B\phi(z)^{r/2}} \le e^{-\beta z^r} \tag{41}$$

If $m+n \leq \phi(z)$, then $s \leq \sqrt{1+\phi(z)}$ and $z-s \geq z-\sqrt{1+\phi(z)} = \sqrt{\beta} z^{r/2}$. By (17), this means that $l_{m,n}(z) \leq \frac{1}{\pi} e^{-\beta z^r}$. So

$$\sum_{m+n \leq \phi(z)} |\rho_{m,n}| l_{m,n}(z) \leq A e^{-\beta z^r} \tag{42}$$

for $A := \frac{1}{\pi} \sum_{m,n} e^{-B(m+n)^{r/2}}$.

On the other hand, using Lemma 3 with $\nu := r/2$, if $\phi(z) \geq z_0$,

$$\begin{aligned}\sum_{m+n \geq \phi(z)} |\rho_{m,n}| l_{m,n}(z) &\leq \frac{4}{\pi B r} \phi(z)^{2-r/2} e^{-B\phi(z)^{r/2}} \\ &\leq \frac{4}{\pi B r} z^{4-r} e^{-\beta z^r} \end{aligned} \tag{43}$$

by (17) and (41). Combining (42) and (43) yields the announced result. The bound on $\widetilde{W_\rho}$ is then a direct consequence of (12).

5.2 Proof of Proposition 2

Let $\phi(z) := \theta z^2 - 1$, where $\theta := \frac{1}{(1+\sqrt{B})^2}$ is the solution in $(0,1)$ of $(1-\sqrt{\theta})^2 = B\theta$.

When $m+n \leq \phi(z)$, then $s \leq \sqrt{\theta} z$ and $z - s \geq z(1-\sqrt{\theta}) = \sqrt{B\theta} z$. By (17), this means that $l_{m,n}(z) \leq \frac{1}{\pi} e^{-B\theta z^2}$. So

$$\sum_{m+n \leq \phi(z)} |\rho_{m,n}| l_{m,n}(z) \leq A e^{-B\theta z^2} \tag{44}$$

for $A := \frac{1}{\pi} \sum_{m,n} e^{-B(m+n)}$.

On the other hand, by Lemma 3, if $\phi(z) \geq z_0$,

$$\begin{aligned}\sum_{m+n \geq \phi(z)} |\rho_{m,n}| l_{m,n}(z) &\leq \frac{2}{\pi B} \phi(z) e^{-B\phi(z)} \\ &\leq \frac{2\theta e^B}{\pi B} z^2 e^{-B\theta z^2} \end{aligned} \tag{45}$$

by (17) and (41). Combining (44) and (45) yields the announced result. The bound on $\widetilde{W_\rho}$ is then a direct consequence of (12).

5.3 Proof of Proposition 3

By (6) the term $b_1^2(n)$ can be bounded as follows

$$b_1^2(n) = \sum_{j+k \geq N} |\rho_{j,k}|^2 \leq \sum_{j+k \geq N} \exp(-2B(j+k)^{r/2}).$$

Compare to the double integral and change to polar coordinates to get

$$b_1^2(n) \quad \leq \quad c_1 N^{2-r/2} \exp(-2BN^{r/2}).$$

5.4 Proof of Proposition 4

To study the term $b_2^2(n)$, we denote

$$\mathcal{F}_1[p_\rho(\cdot|\phi)](t) := E_\rho[e^{itX}|\Phi = \phi] = \widetilde{W}_\rho(t\cos\phi, t\sin\phi),$$

the Fourier transform with respect to the first variable.

$$\begin{aligned}
E[\hat{\rho}^\eta_{j,k}] &= E[G_{j,k}(\frac{Y}{\sqrt{\eta}}, \Phi)] = E[f^{\eta,\delta}_{j,k}(\frac{Y}{\sqrt{\eta}})e^{-i(j-k)\Phi}] \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi} \int f^{\eta,\delta}_{j,k}(y)\sqrt{\eta}p^\eta_\rho(y\sqrt{\eta}|\phi)dyd\phi \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi}\frac{1}{2\pi}\int \widetilde{f}^{\eta,\delta}_{j,k}(t)\mathcal{F}_1[\sqrt{\eta}p^\eta_\rho(\cdot\sqrt{\eta}|\phi)](t)dtd\phi \\
&= \frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi}\frac{1}{2\pi}\int_{|t|\leq 1/\delta} \widetilde{f}_{j,k}(t)e^{\gamma t^2}\mathcal{F}_1[p_\rho(\cdot|\phi)](t)\widetilde{N}^\eta(t)dtd\phi.
\end{aligned}$$

As $\widetilde{N}^\eta(t) = e^{-\gamma t^2}$ and by using the Cauchy-Schwarz inequality we have

$$\begin{aligned}
\left|E[\hat{\rho}^\eta_{j,k}] - \rho_{j,k}\right|^2 &= \left|\frac{1}{\pi}\int_0^\pi e^{-i(j-k)\phi}\frac{1}{2\pi}\int_{|t|>1/\delta}\widetilde{f}_{j,k}(t)\mathcal{F}_1[p_\rho(\cdot|\phi)](t)dtd\phi\right|^2 \\
&\leq \frac{1}{\pi}\int_0^\pi\left(\frac{1}{2\pi}\int_{|t|>1/\delta}\left|\widetilde{f}_{j,k}(t)\widetilde{W}_\rho(t\cos\phi, t\sin\phi)\right|dt\right)^2 d\phi.
\end{aligned}$$

If $1/\delta \geq 2\sqrt{N} \geq 2s$ with $s = \sqrt{j+k+1}$, then whenever $t \geq 1/\delta$ we get by Lemma 2

$$\begin{aligned}
|\widetilde{f}_{j,k}(t)| &= \pi^2|t|l_{j,k}(t/2) \\
&\leq \pi|t|e^{-\frac{1}{4}(|t|-2s)^2}.
\end{aligned}$$

On the other hand, by Propositions 1 and 2 we have

$$|\widetilde{W}_\rho(t\cos\phi, t\sin\phi)| \leq A(\frac{|t|}{2})e^{-\beta(\frac{|t|}{2})^r}$$

for $\beta := \frac{B}{(1+\sqrt{B})^2}$ in the case $r=2$, or for arbitrary $\beta < B$ and t large enough in the case $0 < r < 2$. In both cases A is a polynom of degree $4-r$. We deduce the inequality

$$\begin{aligned}
\left|E[\hat{\rho}^\eta_{j,k}] - \rho_{j,k}\right|^2 &\leq C\left(\int_{\frac{1}{\delta}}^\infty t^{5-r}e^{-\frac{1}{4}(t-2s)^2-\beta 2^{-r}t^r}dt\right)^2 \\
&\leq C(\frac{1}{\delta})^{12-4r}e^{-\frac{1}{2}(\frac{1}{\delta}-2\sqrt{N})^2-\beta 2^{1-r}(\frac{1}{\delta})^r}
\end{aligned}$$

by Lemma 8 in [5], hence

$$b_2(n)^2 \leq CN^2(\frac{1}{\delta})^{12-4r}e^{-\frac{1}{2}(\frac{1}{\delta}-2\sqrt{N})^2-\beta 2^{1-r}(\frac{1}{\delta})^r}$$

which covers both cases in the proposition.

5.5 Proof of Proposition 5

Let us write $\sigma^2_{j,k}(n) := E\left|\hat{\rho}^{\eta}_{j,k} - E[\hat{\rho}^{\eta}_{j,k}]\right|^2$. We bound it by

$$\begin{aligned}
\sigma^2_{j,k}(n) &= E\left|\frac{1}{n}\sum_{\ell=1}^{n}\left(G_{j,k}(\frac{Y_\ell}{\sqrt{\eta}},\Phi_\ell) - E[G_{j,k}(\frac{Y_\ell}{\sqrt{\eta}},\Phi_\ell)]\right)\right|^2 \\
&= \frac{1}{n}E\left|G_{j,k}(\frac{Y}{\sqrt{\eta}},\Phi) - E[G_{j,k}(\frac{Y}{\sqrt{\eta}},\Phi)]\right|^2 \\
&\leq \frac{1}{n}E\left|G_{j,k}(\frac{Y}{\sqrt{\eta}},\Phi)\right|^2. \qquad (46)
\end{aligned}$$

Proof of (31) For $0 < \eta \leq 1/2$, let us denote by K_δ the function with the following Fourier transform $\widetilde{K}_\delta(t) = \mathbb{I}(|t| \leq \frac{1}{\delta})e^{\gamma t^2}$, then $\widetilde{f}^{\eta,\delta}_{j,k} = \widetilde{f}_{j,k}(t)\widetilde{K}_\delta(t)$ and we have

$$\begin{aligned}
\sigma^2_{j,k}(n) &\leq \frac{1}{n}E\left|f^{\eta,\delta}_{j,k}(\frac{Y}{\sqrt{\eta}})e^{-i(j-k)\Phi}\right|^2 \\
&\leq \frac{1}{n}E\left|f_{j,k} * K_\delta(\frac{Y}{\sqrt{\eta}})\right|^2 \\
&\leq \frac{1}{n}E\left|\int f_{j,k}(t)K_\delta(\frac{Y}{\sqrt{\eta}}-t)dt\right|^2.
\end{aligned}$$

By using the Cauchy-Schwarz inequality

$$\begin{aligned}
\sigma^2_{j,k}(n) &\leq \frac{1}{n}\int |f_{j,k}(t)|^2 dt E\int\left|K_\delta(\frac{Y}{\sqrt{\eta}}-t)\right|^2 dt \\
&\leq \frac{1}{n}\int |f_{j,k}(t)|^2 dt E\frac{1}{2\pi}\int\left|\widetilde{K}_\delta(u)e^{-iu\frac{Y}{\sqrt{\eta}}}\right|^2 du \\
&\leq \frac{1}{n\pi}\|f_{j,k}\|_2^2\int_0^{1/\delta} e^{2\gamma u^2}du.
\end{aligned}$$

Then,

$$\sigma^2(n) \leq \frac{C}{n\pi}\sum_{j+k=0}^{N-1}\|f_{j,k}\|_2^2\frac{\eta\delta}{1-\eta}e^{\frac{2\gamma}{\delta^2}}.$$

By Lemma 4 we have $\sum_{j+k=0}^{N-1} \|f_{j,k}\|_2^2 \leq C_2 N^{17/6}$ thus

$$\sigma^2(n) \leq \frac{C_1 \eta \delta N^{17/6}}{n\pi(1-\eta)} e^{\frac{2\gamma}{\delta^2}}.$$

Proof of (32) and (33) By (28), for $1/2 < \eta \leq 1$,

$$\begin{aligned}
\sigma_{j,k}^2(n) &\leq \frac{1}{n} E\left| f_{j,k}^{\eta}(\frac{Y}{\sqrt{\eta}}) e^{-i(j-k)\Phi} \right|^2 \\
&\leq \frac{1}{n\pi} \int_0^{\pi} \int \left| f_{j,k}^{\eta}(y) \right|^2 \sqrt{\eta} p_{\rho}^{\eta}(\sqrt{\eta} y|\phi) dy d\phi \\
&\leq \frac{1}{n\pi} \left\| f_{j,k}^{\eta} \right\|_{\infty}^2
\end{aligned}$$

For $1/2 < \eta < 1$, by Lemma 5,

$$\sigma^2(n) \leq \frac{C_\infty N^{1/3}}{n\pi} e^{8\gamma N}.$$

For $\eta = 1$, by Lemma 6

$$\begin{aligned}
\sigma_{j,k}^2(n) &\leq \frac{1}{n} \int_0^{\pi} \int |f_{j,k}(x)|^2 p_{\rho}(x,\phi) dx d\phi \\
&\leq \frac{C}{n} \|f_{j,k}\|_2^2
\end{aligned}$$

hence by Lemma 4,

$$\sigma^2(n) \leq C \frac{C_2 N^{17/6}}{n}.$$

5.6 Proof of Proposition 6

It is easy to see that

$$\mathcal{F}\left[E[\widehat{W}_h^{\eta}] \right](w) = \widetilde{W}_{\rho}(w) I(\|w\| \leq 1/h).$$

We have, for n large enough $s_n \geq z_0$ and by (13)

$$\begin{aligned}
\|W_\rho\|_{\overline{D}(s_n)}^2 &\leq C(B,r) \int_{\|z\|>s_n} \|z\|^{8-2r} \exp(-2\beta \|z\|^r) dz \\
&\leq C(B,r) \int_0^{2\pi} \int_{s_n}^{\infty} t^{9-2r} \exp(-2\beta t^r) dt d\phi \\
&\leq C_1 s_n^{10-3r} e^{-2\beta s_n^r},
\end{aligned}$$

where $\beta < B$ and for n large enough in the case $0 < r < 2$, respectively $\beta = B/(1+\sqrt{B})^2$ in the case $r = 2$. Now we write for the $\mathbb{L}_2$ bias of our estimator :

$$\begin{aligned}
\|E[\widehat{W}_h^\eta] - W_\rho\|_{D(s_n)}^2 &\leq \|E[\widehat{W}_h^\eta] - W_\rho\|_2^2 = \frac{1}{(2\pi)^2}\|\mathcal{F}\left[E[\widehat{W}_h^\eta]\right] - \widetilde{W}_\rho\|_2^2 \\
&= \frac{1}{(2\pi)^2}\int \left|\widetilde{W}_\rho(w)\right|^2 I(\|w\| > 1/h)\, dw \\
&\leq \frac{C^2(B,r)}{(2\pi)^2}\int_{\|w\|>1/h} \|w\|^{2(4-r)} e^{-2^{1-r}\beta\|w\|^r}\, dw \\
&\leq C_2 h^{3r-10} e^{-\frac{2^{1-r}\beta}{h^r}},
\end{aligned}$$

by the assumption on our class and (14), for $0 < r < 2$. The case $r = 2$ is similar.

As for the variance of our estimator :

$$\begin{aligned}
V\left[\widehat{W}_h^\eta\right] &= E\left[\left\|\widehat{W}_h^\eta - E\left[\widehat{W}_h^\eta\right]\right\|_{D(s_n)}^2\right] \\
&= \frac{1}{\pi^2 n}\left\{E\left[\left\|K_h^\eta\left([\cdot,\Phi] - \frac{Y}{\sqrt{\eta}}\right)\right\|_{D(s_n)}^2\right]\right. \\
&\quad \left. - \left\|E\left[K_h^n\left([\cdot,\Phi] - \frac{Y}{\sqrt{\eta}}\right)\right]\right\|_{D(s_n)}^2\right\}. \qquad (47)
\end{aligned}$$

On the one hand, by using two-dimensional Plancherel formula and the Fourier transform shown above, we get :

$$\left\|E\left[K_h^n\left([\cdot,\Phi] - \frac{Y}{\sqrt{\eta}}\right)\right]\right\|_{D(s_n)}^2 \leq \pi^2\int |W_\rho(w)|^2 dw \leq \pi^2. \qquad (48)$$

In the last inequality we have used the fact that $\|W_\rho\|_2^2 = \mathrm{Tr}(\rho^2) \leq 1$ where ρ is the density matrix corresponding to the Wigner function W_ρ. On the other hand, the dominant term in the variance will be given by

$$\begin{aligned}
& E\left[\left\|K_h^\eta\left([\cdot,\Phi] - \frac{Y}{\sqrt{\eta}}\right)\right\|_{D(s_n)}^2\right] \\
= & \int_0^\pi \int\int_{D(s_n)} \left(K_h^\eta([z,\phi] - y/\sqrt{\eta})\right)^2 dz\, p_\rho^\eta(y,\phi) dy d\phi \\
= & \int_0^\pi \int_{D(s_n)}\int \left(K_h^\eta(u)\right)^2 \sqrt{\eta} p_\rho^\eta(([z,\phi] - u)\sqrt{\eta},\phi) du dz d\phi \\
= & \int \left(K_h^\eta(u)\right)^2 \int_{D(s_n)}\int_0^\pi p_\rho(\cdot,\phi) * NN^\eta([z,\phi] - u) d\phi dz du \\
\leq & M(\eta)\pi s_n^2 \int (K_h^\eta(u))^2 du,
\end{aligned}$$

using Lemma 6 below and the constant $M(\eta) > 0$ depending only on η, defined therein. Indeed, let us note that $\sqrt{\eta} p_\rho^\eta(\cdot \sqrt{\eta}, \phi)$ is the density of $Y/\sqrt{\eta} = X + \sqrt{(1-\eta)/(2\eta)}\varepsilon$ and let us call NN^η the Gaussian density of the noise as normalized in this last equation.

Let us first compute, by Plancherel formula, $\|K_h^\eta\|_2^2$ and get

$$\begin{aligned}
\|K_h^\eta\|_2^2 &= \frac{1}{2\pi} \int |\widetilde{K}_h^\eta(t)|^2 dt = \frac{1}{2\pi} \int_{|t| \le 1/h} \frac{t^2}{4\widetilde{N}^2(t\sqrt{(1-\eta)/(2\eta)})} dt \\
&= \frac{1}{4\pi} \int_0^{1/h} t^2 \exp\left(t^2 \frac{1-\eta}{2\eta}\right) dt \\
&= \frac{1}{4\pi h} \frac{\eta}{1-\eta} \exp\left(\frac{1-\eta}{2\eta h^2}\right)(1+o(1)), \text{ as } h \to 0.
\end{aligned}$$

We replace in the second order moment, then as $h \to 0$

$$E\left[\left\| K_h^\eta\left([\cdot, \Phi] - \frac{Y}{\sqrt{\eta}}\right)\right\|_{D(s_n)}^2\right] \le \frac{M(\eta) s_n^2}{16\gamma h} \exp\left(\frac{2\gamma}{h^2}\right)(1+o(1)). \tag{49}$$

The result about the variance of the estimator is obtained from (47)-(49).

Lemma 6. *For every $\rho \in \mathcal{R}(B, r)$ and $0 < \eta < 1$, we have that the corresponding probability density p_ρ satisfies*

$$\begin{aligned}
0 &\le \int_0^\pi p_\rho(\cdot, \phi) * NN^\eta(x) d\phi \le M(\eta), \\
0 &\le \int_0^\pi p_\rho(x, \phi) d\phi \le C
\end{aligned}$$

for all $x \in \mathbb{R}$ eventually depending on ϕ, where $M(\eta) > 0$ is a constant depending only on fixed η and $C > 0$.

Démonstration. Indeed, using inverse Fourier transform and the fact that $\left|\widetilde{W}_\rho(w)\right| \le 1$ we get :

$$\begin{aligned}
& \left| \int_0^\pi p_\rho(\cdot, \phi) * NN^\eta(x) d\phi \right| \\
\le & \left| \int_0^\pi \frac{1}{2\pi} \int e^{-itx} \mathcal{F}_1[p_\rho(\cdot, \phi)](t) \cdot \widetilde{NN}^\eta(t) dt d\phi \right| \\
\le & \ c(\eta) \int_0^\pi \int \left|\widetilde{W}_\rho(t\cos\phi, t\sin\phi)\right| \exp\left(-\frac{t^2(1-\eta)}{4\eta}\right) dt d\phi \\
\le & \ c(\eta) \int \frac{1}{\|w\|} \left|\widetilde{W}_\rho(w)\right| \exp\left(-\frac{\|w\|^2(1-\eta)}{4\eta}\right) dw \le M(\eta),
\end{aligned}$$

where $c(\eta)$, $M(\eta)$ are positive constants depending only on $\eta \in (0, 1)$. $\square$

Références

[1] L. Artiles, R. Gill, and M. Guţă. An invitation to quantum tomography. *J. Royal Statist. Soc. B (Methodological)*, 67 :109–134, 2005.

[2] J. M. Aubry. Ultrarapidly decreasing ultradifferentiable functions, Wigner distributions and density matrices. Submitted to J. London Math. Soc., 2008.

[3] K. Banaszek, G. M. D'Ariano, M. G. A. Paris, and M. F. Sacchi. Maximum-likelihood estimation of the density matrix. *Physical Review A*, 61(R10304), 2000.

[4] C. Butucea, M. Guţă, and L. Artiles. Minimax and adaptive estimation of the Wigner function in quantum homodyne tomography with noisy data. *Ann. Statist.*, 35(2) :465–494, 2007.

[5] C. Butucea and A. B. Tsybakov. Sharp optimality for density deconvolution with dominating bias. i and ii. *Theory Probab. Appl.*, 2007.

[6] L. Cavalier. Efficient estimation of a density in a problem of tomography. *Ann. Statist.*, 28 :630–647, 2000.

[7] G. M. D'Ariano. Tomographic measurement of the density matrix of the radiation field. *Quantum Semiclass. Optics*, 7 :693–704, 1995.

[8] G. M. D'Ariano. Tomographic methods for universal estimation in quantum optics. In *International School of Physics Enrico Fermi*, volume 148. IOS Press, 2002.

[9] G. M. D'Ariano, U. Leonhardt, and H. Paul. Homodyne detection of the density matrix of the radiation field. *Phys. Rev. A*, 52 :R1801–R1804, 1995.

[10] G. M. D'Ariano, C. Macchiavello, and M. G. A. Paris. Detection of the density matrix through optical homodyne tomography without filtered back projection. *Phys. Rev. A*, 50 :4298–4302, 1994.

[11] G. M. D'Ariano and M. G. A. Paris. Adaptive quantum tomography. *Physical Review A*, 60(518), 1999.

[12] G.M. D'Ariano, L. Maccone, and M. F. Sacchi. Homodyne tomography and the reconstruction of quantum states of light, 2005.

[13] A. Goldenshluger and V. Spokoiny. On the shape-from-moments problem and recov-

ering edges from noisy Radon data. *Probab. Theory Related Fields*, 128(1) :123–140, 2004.

[14] M. Guţă and L. Artiles. Minimax estimation of the Wigner in quantum homodyne tomography with ideal detectors. *Math. Methods Statist.*, 16(1) :1–15, 2007.

[15] M. I. Guţă. Maximum likelihood estimation of the density matrix through quantum tomography. Manuscript, 2007.

[16] A. P. Korostelëv and A. B. Tsybakov. *Minimax theory of image reconstruction*, volume 82 of *Lecture Notes in Statistics*. Springer-Verlag, New York, 1993.

[17] I. Krasikov. Inequalities for Laguerre polynomials. *East J. Approx.*, 11(3) :257–268, 2005.

[18] I. Krasikov. Inequalities for orthonormal Laguerre polynomials. *J. Approx. Theory*, 144(1) :1–26, 2007.

[19] U. Leonhardt. *Measuring the Quantum State of Light.* Cambridge University Press, 1997.

[20] U. Leonhardt, H. Paul, and G. M. D'Ariano. Tomographic reconstruction of the density matrix via pattern functions. *Phys. Rev. A*, 52 :4899–4907, 1995.

[21] Johnstone Iain M. and Silverman Bernard W. Speed of estimation in positron emission tomography and related inverse problems. *Ann. Statist.*, 18(1) :251–280, 1990.

[22] K. Meziani. Nonparametric estimation of the purity of a quantum state in quantum homodyne tomography with noisy data. *Mathematical Methods of Statistics*, 16(4) :1–15, 2007.

[23] T. Richter. Pattern functions used in tomographic reconstruction of photon statistics revisited. *Phys. Lett. A*, 211 :327–330, 1996.

[24] T. Richter. Realistic pattern functions for optical homodyne tomography and determination of specific expectation values. *Physical Review A*, 61(063819), 2000.

[25] D. J. Smithey, M. Beck, M. J. Cooper, and M. G. Raymer. Experimental determination of number-phase uncertainty relations. *Optics Letters*, 18 :1259–1261, 1993.

[26] D. T. Smithey, M. Beck, M. G. Raymer, and A. Faridani. Measurement of the Wigner distribution and the density matrix of a light mode using optical homodyne tomogra-

phy : Application to squeezed states and the vacuum. *Phys. Rev. Lett.*, 70 :1244–1247, 1993.

[27] G. Szegö. *Orthogonal polynomials.* American Mathematical Society Colloquium Publications, Vol. 23. Revised ed. American Mathematical Society, Providence, R.I., 1959.

[28] Y. Vardi, L. A. Shepp, and L. Kaufman. A statistical model for positron emission tomography. *J. Am. Stat. Assoc.*, 80 :8–37, 1985.

[29] K. Vogel and H. Risken. Determination of quasiprobability distributions in terms of probability distributions for the rotated quadrature phase. *Phys. Rev. A*, 40 :2847–2849, 1989.